Automatic Genre Classification: A Study of the Viability of High-Level Features for Music Classification

Cory McKay
Music Technology Area, McGill University
cory.mckay@mail.mcgill.ca

Goals

- Demonstrate power of high-level musical features by successfully using them to classify symbolic recordings by genre
- Show effectiveness of large feature libraries when combined with effective feature selection
- Demonstrate that automatic genre classification has potential with realistically large taxonomies

Methodology

- Compiled a catalogue of 160 high-level features
- Collected 950 MIDI recordings for training and testing
- Used sophisticated combination of neural networks and k-nearest neighbour classifiers
- Performed feature selection and weighting with genetic algorithms
- Developed an easy to use and flexible GUI

Results

- 86% success with 9 leaf categories
- 57% success with 38 leaf categories

Applications of automatic genre classification

- Automatic classification of large dynamic on-line music databases
- Aid to musicological and psychological research on genre
- Techniques can be applied to other types of classification:
    - Performer or composer style, geographical, cultural or temporal influences, mood, character, personal preferences

Features

Overview

- Catalogue of 160 high level features compiled (111 implemented)
  - Applicable to types of research other than genre classification
- Acquired information on which features were most important in which taxonomical contexts using automatic feature weighting
- Avoided features dependant on specialized theoretical models

Reasons for using high-level musical features recordings rather than features based on signal processing

- Have relevance to musicological and analytical studies of genre
- Take advantage of abstractions that have proven usefulness to musical experts (e.g. musicians, theorists and composers)
- Can use optical music recognition to classify scores when no audio recordings available
- Research can be adapted to audio classification as automatic transcription techniques improve

Types of features implemented

- **Instrumentation:** Which instruments were present and which ones were given particular importance. Relative importance of pitched and non-pitched instruments.
- **Texture:** Number of independent voices, their interactions and the relative importance of different voices and registers.
- **Rhythm:** Time intervals between note attacks and durations of notes. Meter, use of rubato and other rhythmic variations.
- **Dynamics:** Average loudness and variability in dynamics.
- **Pitch Statistics:** Occurrence rates of different notes, range, degree of tonality and variety of pitches.
- **Melody:** Frequencies and variability of melodic intervals, types of melodic contours and types of phrases.
- **Chords:** Types and frequencies of vertical intervals and vertical note densities.

Genre Taxonomies

Overview

- Used a hierarchical taxonomy with two important modifications:
  - A recording could belong to more than one category
  - A category could be a child of more than one parent
- Based on All Music Guide, retailers, magazines and fan sites

Small taxonomy for comparing this system to existing systems:

Jazz
- Bebop
- Jazz Soul
- Swing

Popular
- Rap
- Punk
- Country

Western Classical
- Baroque
- Modern Classical
- Romantic

Large taxonomy for testing system under realistic conditions:

Country
- Bluegrass
- Contemporary
- Trad. Country

Jazz
- Bop
  - Bebop
  - Cool
- Fusion
  - Bossa Nova
  - Jazz Soul
  - Smooth Jazz
- Ragtime
- Swing

Modern Pop
- Adult Contemp.
- Dance
  - Dance Pop
  - Pop Rap
  - Techno
- Smooth Jazz

Rap
- Hardcore Rap
- Pop Rap

Rhythm and Blues
- Blues
  - Blues Rock
  - Chicago Blues
  - Country Blues
  - Soul Blues
- Funk
- Jazz Soul
- Rock and Roll
- Soul

Rock
- Classic Rock
  - Blues Rock
  - Hard Rock
  - Psychedelic
- Modern Rock
  - Alt. Rock
  - Hard Rock
  - Metal
  - Punk

Western Classical
- Baroque
- Classical
- Early Music
  - Medieval
  - Renaissance
- Modern Classical
- Romantic

Western Folk
- Bluegrass
- Celtic
- Country Blues
- Flamenco

Worldbeat
- Latin
  - Bossa Nova
  - Salsa
  - Tango
- Reggae

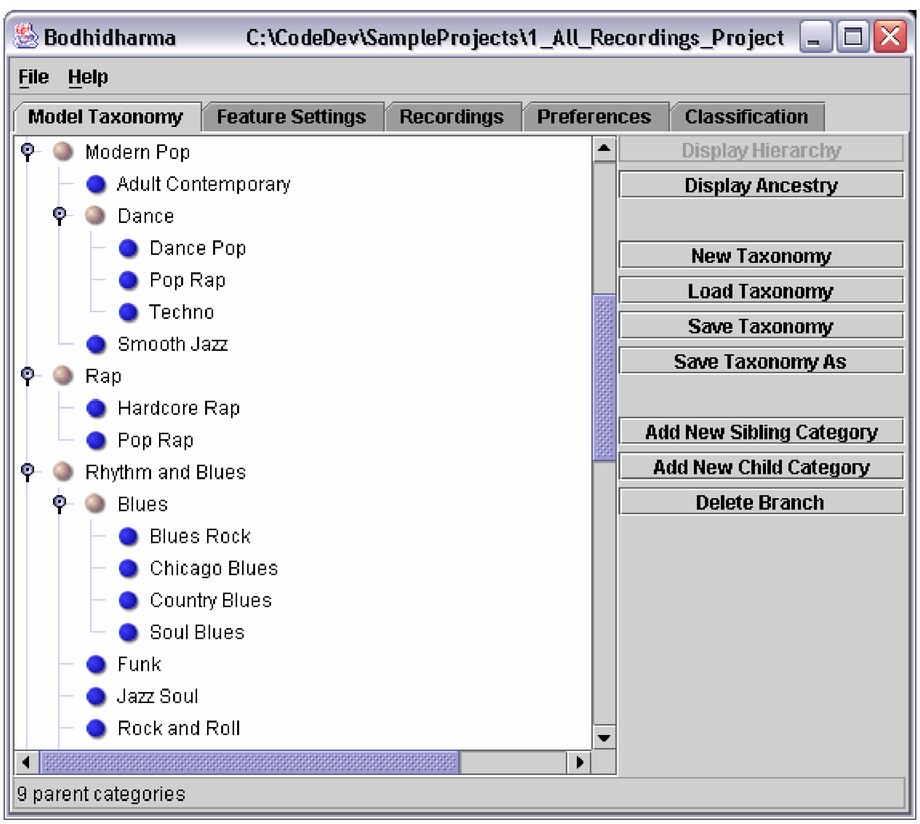
Software System

Classification system

- Used two types of classifiers:
  - K-Nearest Neighbour
    - Fast to train
  - Feedforward Neural Networks
    - Can model sophisticated relationships between features
- Used a combination of three classification methodologies:
  - Hierarchical
  - Round Robin
  - Flat
- Feature selection and weighting performed using genetic algorithms
  - Made it possible to make specialized classifiers using only features ideally suited to particular classification contexts

Interface

- Designed to be easy to use for people with little technical background
- Highly customizable
  - Configurations stored in XML files grouped into projects

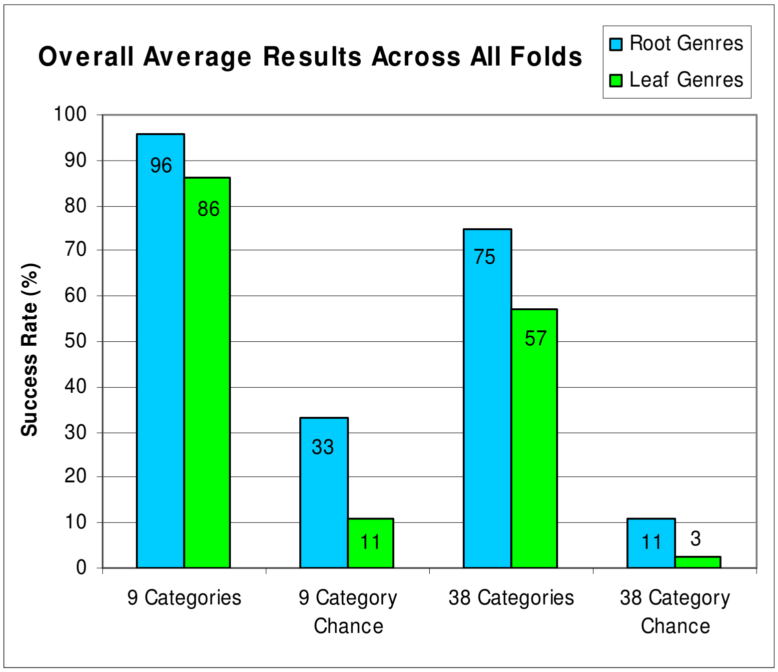


Experimental Results

Experiment

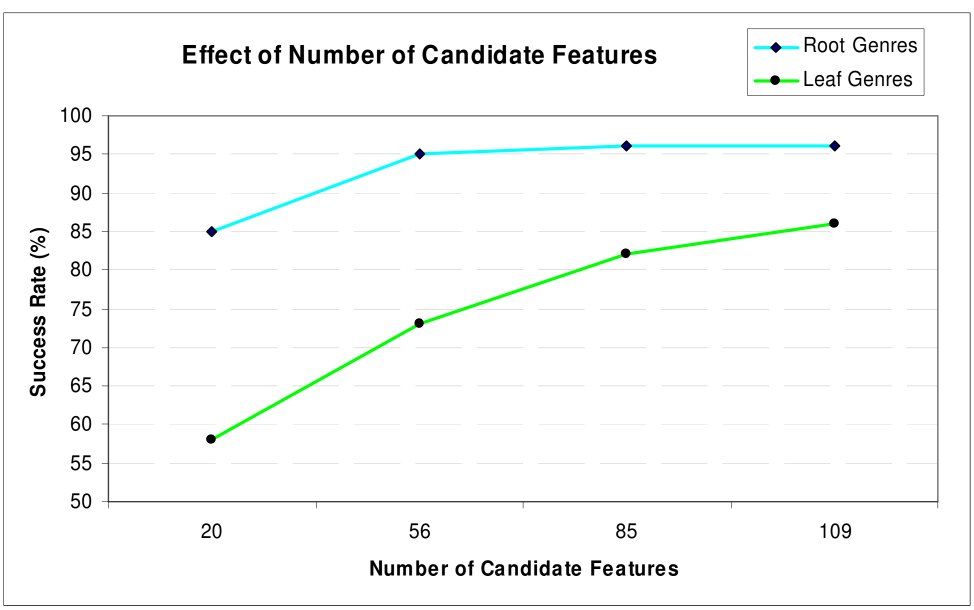
- 950 MIDI files for training and testing
- 5 fold cross-validation

Classification results



Investigation of importance of a using a large feature library

- Examined effect on success rate of only providing subsets of feature catalogue as candidates to feature selection system



Conclusions

Effectiveness of system

- Success rates with 9 category taxonomy:
  - 86% among 9 leaf categories
    - 96% among the 3 root categories
  - Much better success rates than previous systems dealing with symbolic recordings
  - Comparable with the best results involving audio recordings
- Success rates with 38 category taxonomy:
  - 57% among 38 leaf categories
    - 75% among the 9 root categories
  - No previous research has attempted to classify recordings using such a difficult taxonomy
- Results particularly encouraging considering dealt with added complication that recordings could belong to an unknown number of categories

Additional conclusions

- Effectiveness of high-level features clearly demonstrated
- A large feature library combined with good feature selection improves results
- Not yet at a point where can effectively deal with large realistic taxonomies, but are approaching that point

McGill

Social Sciences and Humanities Research Council of Canada

CIRMMT Centre for Interdisciplinary Research in Music Media and Technology

Fonds de recherche sur la société et la culture Québec